wwPDB X-ray Structure Validation Summary Report (i)

Oct 24, 2024 – 12:45 AM EDT

PDB ID	:	1ZMB
Title	:	Crystal Structure of the Putative Acetylxylan Esterase from Clostridium acetobutylicum, Northeast Structural Genomics Target CaR6
Authors	:	Forouhar, F.; Vorobiev, S.M.; Abashidze, M.; Ciao, M.; Acton, T.B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2005-05-10
Resolution	:	2.61 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

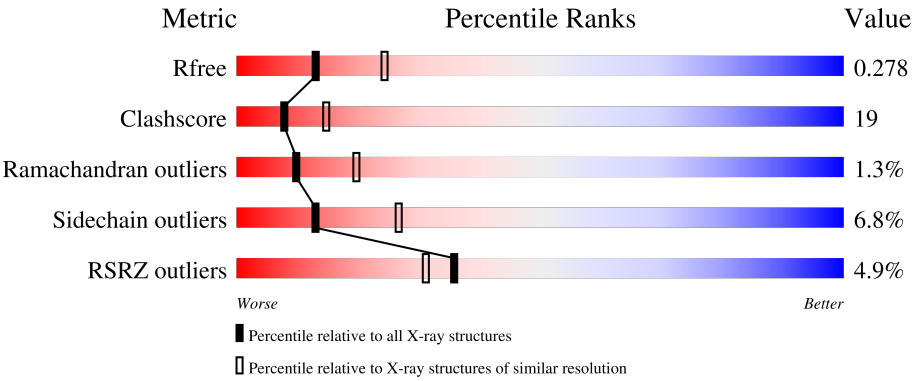
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.61 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



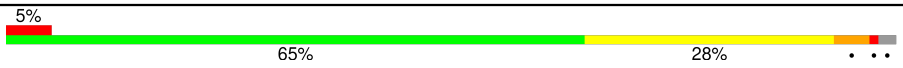
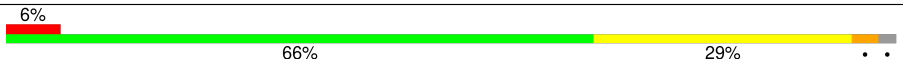
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	164625	4623 (2.64-2.60)
Clashscore	180529	5071 (2.64-2.60)
Ramachandran outliers	177936	5006 (2.64-2.60)
Sidechain outliers	177891	5006 (2.64-2.60)
RSRZ outliers	164620	4622 (2.64-2.60)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	290	2% 68% 27%
1	B	290	2% 68% 27%
1	C	290	8% 63% 31%
1	D	290	4% 65% 29%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	E	290	
1	F	290	

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 13791 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Acetylxylan esterase related enzyme.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	284	Total	C	N	O	S	Se	0	0	0
			2278	1457	381	426	4	10			
1	B	284	Total	C	N	O	S	Se	0	0	0
			2278	1457	381	426	4	10			
1	C	284	Total	C	N	O	S	Se	0	0	0
			2278	1457	381	426	4	10			
1	D	284	Total	C	N	O	S	Se	0	0	0
			2278	1457	381	426	4	10			
1	E	284	Total	C	N	O	S	Se	0	0	0
			2278	1457	381	426	4	10			
1	F	284	Total	C	N	O	S	Se	0	0	0
			2278	1457	381	426	4	10			

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MSE	MET	modified residue	UNP Q97LM8
A	7	MSE	MET	modified residue	UNP Q97LM8
A	13	MSE	MET	modified residue	UNP Q97LM8
A	24	MSE	MET	modified residue	UNP Q97LM8
A	32	MSE	MET	modified residue	UNP Q97LM8
A	40	MSE	MET	modified residue	UNP Q97LM8
A	41	MSE	MET	modified residue	UNP Q97LM8
A	106	MSE	MET	modified residue	UNP Q97LM8
A	260	MSE	MET	modified residue	UNP Q97LM8
A	277	MSE	MET	modified residue	UNP Q97LM8
A	283	LEU	-	cloning artifact	UNP Q97LM8
A	284	GLU	-	cloning artifact	UNP Q97LM8
A	285	HIS	-	cloning artifact	UNP Q97LM8
A	286	HIS	-	cloning artifact	UNP Q97LM8
A	287	HIS	-	cloning artifact	UNP Q97LM8
A	288	HIS	-	cloning artifact	UNP Q97LM8
A	289	HIS	-	cloning artifact	UNP Q97LM8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	290	HIS	-	cloning artifact	UNP Q97LM8
B	1	MSE	MET	modified residue	UNP Q97LM8
B	7	MSE	MET	modified residue	UNP Q97LM8
B	13	MSE	MET	modified residue	UNP Q97LM8
B	24	MSE	MET	modified residue	UNP Q97LM8
B	32	MSE	MET	modified residue	UNP Q97LM8
B	40	MSE	MET	modified residue	UNP Q97LM8
B	41	MSE	MET	modified residue	UNP Q97LM8
B	106	MSE	MET	modified residue	UNP Q97LM8
B	260	MSE	MET	modified residue	UNP Q97LM8
B	277	MSE	MET	modified residue	UNP Q97LM8
B	283	LEU	-	cloning artifact	UNP Q97LM8
B	284	GLU	-	cloning artifact	UNP Q97LM8
B	285	HIS	-	cloning artifact	UNP Q97LM8
B	286	HIS	-	cloning artifact	UNP Q97LM8
B	287	HIS	-	cloning artifact	UNP Q97LM8
B	288	HIS	-	cloning artifact	UNP Q97LM8
B	289	HIS	-	cloning artifact	UNP Q97LM8
B	290	HIS	-	cloning artifact	UNP Q97LM8
C	1	MSE	MET	modified residue	UNP Q97LM8
C	7	MSE	MET	modified residue	UNP Q97LM8
C	13	MSE	MET	modified residue	UNP Q97LM8
C	24	MSE	MET	modified residue	UNP Q97LM8
C	32	MSE	MET	modified residue	UNP Q97LM8
C	40	MSE	MET	modified residue	UNP Q97LM8
C	41	MSE	MET	modified residue	UNP Q97LM8
C	106	MSE	MET	modified residue	UNP Q97LM8
C	260	MSE	MET	modified residue	UNP Q97LM8
C	277	MSE	MET	modified residue	UNP Q97LM8
C	283	LEU	-	cloning artifact	UNP Q97LM8
C	284	GLU	-	cloning artifact	UNP Q97LM8
C	285	HIS	-	cloning artifact	UNP Q97LM8
C	286	HIS	-	cloning artifact	UNP Q97LM8
C	287	HIS	-	cloning artifact	UNP Q97LM8
C	288	HIS	-	cloning artifact	UNP Q97LM8
C	289	HIS	-	cloning artifact	UNP Q97LM8
C	290	HIS	-	cloning artifact	UNP Q97LM8
D	1	MSE	MET	modified residue	UNP Q97LM8
D	7	MSE	MET	modified residue	UNP Q97LM8
D	13	MSE	MET	modified residue	UNP Q97LM8
D	24	MSE	MET	modified residue	UNP Q97LM8
D	32	MSE	MET	modified residue	UNP Q97LM8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	40	MSE	MET	modified residue	UNP Q97LM8
D	41	MSE	MET	modified residue	UNP Q97LM8
D	106	MSE	MET	modified residue	UNP Q97LM8
D	260	MSE	MET	modified residue	UNP Q97LM8
D	277	MSE	MET	modified residue	UNP Q97LM8
D	283	LEU	-	cloning artifact	UNP Q97LM8
D	284	GLU	-	cloning artifact	UNP Q97LM8
D	285	HIS	-	cloning artifact	UNP Q97LM8
D	286	HIS	-	cloning artifact	UNP Q97LM8
D	287	HIS	-	cloning artifact	UNP Q97LM8
D	288	HIS	-	cloning artifact	UNP Q97LM8
D	289	HIS	-	cloning artifact	UNP Q97LM8
D	290	HIS	-	cloning artifact	UNP Q97LM8
E	1	MSE	MET	modified residue	UNP Q97LM8
E	7	MSE	MET	modified residue	UNP Q97LM8
E	13	MSE	MET	modified residue	UNP Q97LM8
E	24	MSE	MET	modified residue	UNP Q97LM8
E	32	MSE	MET	modified residue	UNP Q97LM8
E	40	MSE	MET	modified residue	UNP Q97LM8
E	41	MSE	MET	modified residue	UNP Q97LM8
E	106	MSE	MET	modified residue	UNP Q97LM8
E	260	MSE	MET	modified residue	UNP Q97LM8
E	277	MSE	MET	modified residue	UNP Q97LM8
E	283	LEU	-	cloning artifact	UNP Q97LM8
E	284	GLU	-	cloning artifact	UNP Q97LM8
E	285	HIS	-	cloning artifact	UNP Q97LM8
E	286	HIS	-	cloning artifact	UNP Q97LM8
E	287	HIS	-	cloning artifact	UNP Q97LM8
E	288	HIS	-	cloning artifact	UNP Q97LM8
E	289	HIS	-	cloning artifact	UNP Q97LM8
E	290	HIS	-	cloning artifact	UNP Q97LM8
F	1	MSE	MET	modified residue	UNP Q97LM8
F	7	MSE	MET	modified residue	UNP Q97LM8
F	13	MSE	MET	modified residue	UNP Q97LM8
F	24	MSE	MET	modified residue	UNP Q97LM8
F	32	MSE	MET	modified residue	UNP Q97LM8
F	40	MSE	MET	modified residue	UNP Q97LM8
F	41	MSE	MET	modified residue	UNP Q97LM8
F	106	MSE	MET	modified residue	UNP Q97LM8
F	260	MSE	MET	modified residue	UNP Q97LM8
F	277	MSE	MET	modified residue	UNP Q97LM8
F	283	LEU	-	cloning artifact	UNP Q97LM8

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	284	GLU	-	cloning artifact	UNP Q97LM8
F	285	HIS	-	cloning artifact	UNP Q97LM8
F	286	HIS	-	cloning artifact	UNP Q97LM8
F	287	HIS	-	cloning artifact	UNP Q97LM8
F	288	HIS	-	cloning artifact	UNP Q97LM8
F	289	HIS	-	cloning artifact	UNP Q97LM8
F	290	HIS	-	cloning artifact	UNP Q97LM8

- Molecule 2 is water.

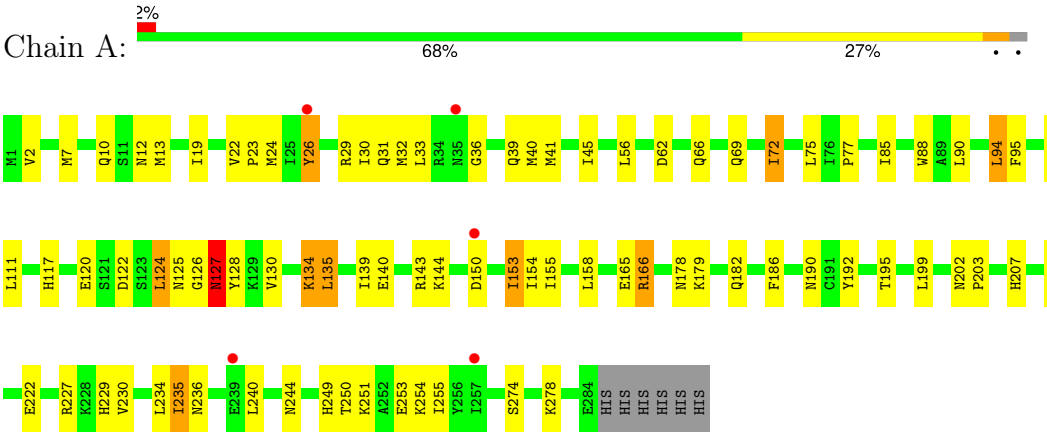
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	26	Total O 26 26	0	0
2	B	27	Total O 27 27	0	0
2	C	21	Total O 21 21	0	0
2	D	12	Total O 12 12	0	0
2	E	19	Total O 19 19	0	0
2	F	18	Total O 18 18	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

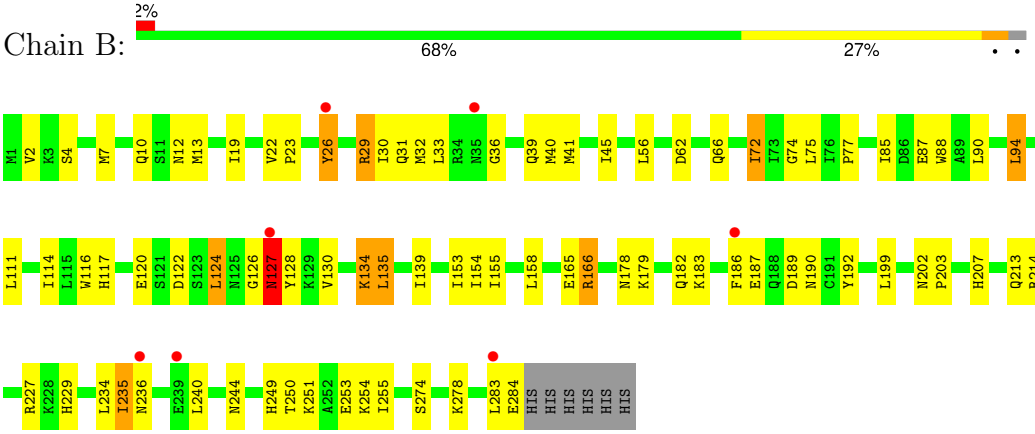
• Molecule 1: Acetylxylan esterase related enzyme

Chain A:



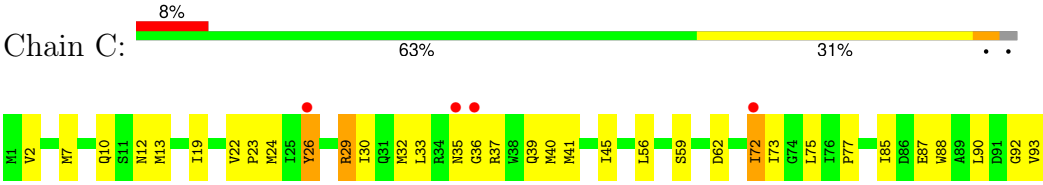
• Molecule 1: Acetylxylan esterase related enzyme

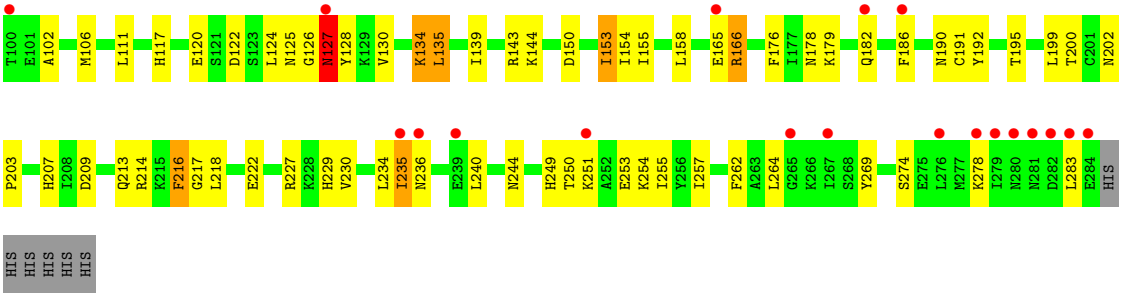
Chain B:



• Molecule 1: Acetylxylan esterase related enzyme

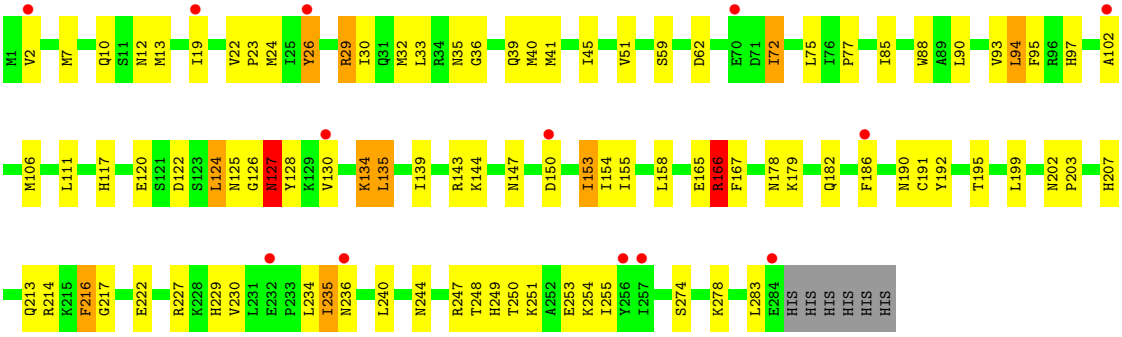
Chain C:





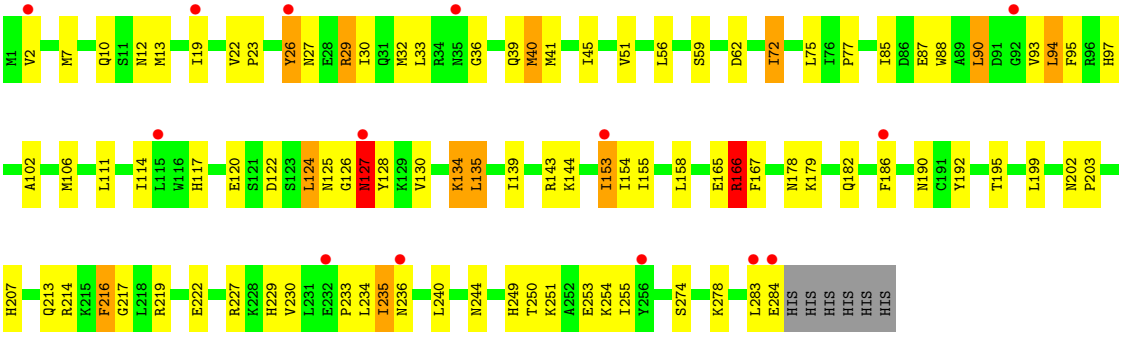
• Molecule 1: Acetylxylan esterase related enzyme

Chain D:



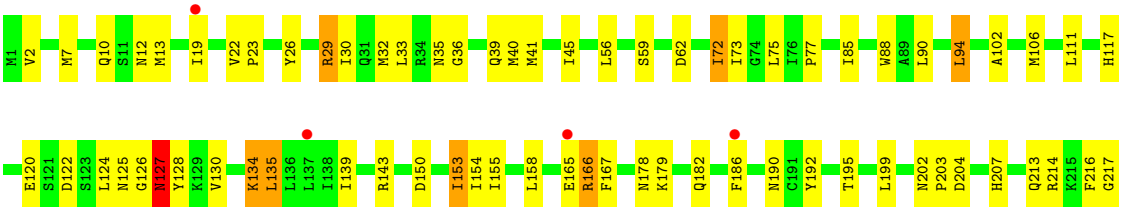
• Molecule 1: Acetylxylan esterase related enzyme

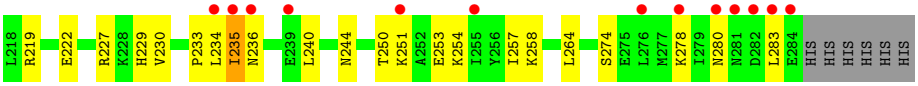
Chain E:



• Molecule 1: Acetylxylan esterase related enzyme

Chain F:




L218
R219
E222
R227
K228
H229
V230
P233
L234
T235
N236
E239
L240
N244
T250
K251
A252
E253
K254
I255
Y256
I257
K258
L264
S274
E275
L276
M277
K278
I279
N280
N281
D282
L283
E284
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	71.00Å 87.94Å 103.51Å 106.50° 100.22° 113.80°	Depositor
Resolution (Å)	29.31 – 2.61 29.31 – 2.61	Depositor EDS
% Data completeness (in resolution range)	82.9 (29.31-2.61) 92.9 (29.31-2.61)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	3.91 (at 2.61Å)	Xtriage
Refinement program	CNS 1.1, XTALVIEW	Depositor
R, R_{free}	0.243 , 0.269 0.255 , 0.278	Depositor DCC
R_{free} test set	5834 reflections (9.67%)	wwPDB-VP
Wilson B-factor (Å^2)	34.8	Xtriage
Anisotropy	0.199	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 39.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.014 for -h,-k,h+k+l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	13791	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.72% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.42	0/2317	0.58	0/3105
1	B	0.42	0/2317	0.57	0/3105
1	C	0.45	0/2317	0.58	0/3105
1	D	0.44	0/2317	0.58	0/3105
1	E	0.42	0/2317	0.57	0/3105
1	F	0.43	0/2317	0.57	0/3105
All	All	0.43	0/13902	0.58	0/18630

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2278	0	2245	84	0
1	B	2278	0	2245	90	0
1	C	2278	0	2245	98	0
1	D	2278	0	2245	91	0
1	E	2278	0	2245	89	0
1	F	2278	0	2245	92	0
2	A	26	0	0	5	0
2	B	27	0	0	1	0
2	C	21	0	0	5	0
2	D	12	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	E	19	0	0	1	0
2	F	18	0	0	2	0
All	All	13791	0	13470	529	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 19.

The worst 5 of 529 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:182:GLN:HG2	1:F:186:PHE:CZ	1.39	1.55
1:F:182:GLN:CG	1:F:186:PHE:CZ	2.28	1.16
1:F:182:GLN:CG	1:F:186:PHE:HZ	1.63	1.11
1:A:106:MSE:HE1	1:A:111:LEU:HD22	1.41	1.01
1:B:106:MSE:HE1	1:B:111:LEU:HD22	1.39	1.01

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	282/290 (97%)	262 (93%)	17 (6%)	3 (1%)	12	24
1	B	282/290 (97%)	261 (93%)	18 (6%)	3 (1%)	12	24
1	C	282/290 (97%)	259 (92%)	19 (7%)	4 (1%)	9	18
1	D	282/290 (97%)	260 (92%)	18 (6%)	4 (1%)	9	18
1	E	282/290 (97%)	261 (93%)	17 (6%)	4 (1%)	9	18
1	F	282/290 (97%)	260 (92%)	18 (6%)	4 (1%)	9	18
All	All	1692/1740 (97%)	1563 (92%)	107 (6%)	22 (1%)	10	20

5 of 22 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	166	ARG
1	A	166	ARG
1	B	166	ARG
1	C	283	LEU
1	D	166	ARG

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	244/240 (102%)	227 (93%)	17 (7%)	12	25
1	B	244/240 (102%)	228 (93%)	16 (7%)	14	28
1	C	244/240 (102%)	228 (93%)	16 (7%)	14	28
1	D	244/240 (102%)	226 (93%)	18 (7%)	11	23
1	E	244/240 (102%)	226 (93%)	18 (7%)	11	23
1	F	244/240 (102%)	229 (94%)	15 (6%)	15	32
All	All	1464/1440 (102%)	1364 (93%)	100 (7%)	13	27

5 of 100 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	134	LYS
1	E	72	ILE
1	F	235	ILE
1	D	147	ASN
1	D	216	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 59 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	238	ASN
1	F	182	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	178	ASN
1	F	178	ASN
1	F	31	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	274/290 (94%)	0.05	5 (1%)	67	63	10, 25, 42, 60	0
1	B	274/290 (94%)	0.13	7 (2%)	57	52	10, 26, 40, 57	0
1	C	274/290 (94%)	0.56	24 (8%)	17	15	13, 34, 52, 76	0
1	D	274/290 (94%)	0.58	13 (4%)	37	33	18, 41, 52, 60	0
1	E	274/290 (94%)	0.60	14 (5%)	34	30	18, 38, 53, 64	0
1	F	274/290 (94%)	0.43	17 (6%)	28	23	17, 34, 51, 77	0
All	All	1644/1740 (94%)	0.39	80 (4%)	36	31	10, 33, 51, 77	0

The worst 5 of 80 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	283	LEU	6.3
1	C	283	LEU	6.1
1	D	186	PHE	6.1
1	F	284	GLU	4.5
1	F	137	LEU	4.1

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.